

# Abelian Ramsey Length and Asymptotic Lower Bounds\*

Vincent Jugé†

This technical note aims at evaluating an asymptotic lower bound on *abelian Ramsey lengths* obtained by Tao in [1]. We first provide the minimal amount of background necessary to define abelian Ramsey lengths, and indicate the lower bound of Tao. We then focus on evaluating this lower bound.

## 1 Introduction

Let  $\mathcal{A}$  and  $\mathcal{V}$  be two alphabets. A word on  $\mathcal{A}$  is a finite sequence  $\mathbf{a} = a_1 a_2 \cdots a_k$  of elements of  $\mathcal{A}$ . The elements  $a_i$  are called the letters of the word  $\mathbf{a}$ , and the integer  $k$  is the length of  $\mathbf{a}$ . For all elements  $\alpha \in \mathcal{A}$ , we denote by  $|\mathbf{a}|_\alpha$  the cardinality of the set  $\{i : a_i = \alpha\}$ , i.e. number of occurrences of the letter  $\alpha$  in the word  $\mathbf{a}$ . We also denote by  $\mathcal{A}^*$  the set of all words on  $\mathcal{A}$ . The words  $a_i a_{i+1} \cdots a_j$  with  $1 \leq i \leq j \leq k$ , as well as the empty word, are called factors of  $\mathbf{a}$ .

Consider now a word  $\mathbf{a} = a_1 a_2 \cdots a_k$  in  $\mathcal{A}^*$  and a word  $\mathbf{p} = p_1 p_2 \cdots p_\ell$  in  $\mathcal{V}^*$ . We say that  $\mathbf{a}$  contains  $\mathbf{p}$  in the abelian sense if there exist non-empty words  $\pi_1, \pi_2, \dots, \pi_\ell$  in  $\mathcal{A}^*$  such that the concatenated word  $\pi_1 \pi_2 \cdots \pi_\ell$  is a factor of  $\mathbf{a}$ , and such that, for all integers  $i, j$  and all letters  $\alpha \in \mathcal{A}$ , if  $p_i = p_j$ , then  $|\pi_i|_\alpha = |\pi_j|_\alpha$ . For instance, the word *programmable* contains the word *aab* in the abelian sense, as can be seen by considering the words  $\pi_1 = am$ ,  $\pi_2 = ma$  and  $\pi_3 = ble$ .

From this point on, we consider the infinite alphabet  $\mathcal{V} = \{v_i : i \in \mathbb{N}\}$ , where  $\mathbb{N}$  is the set of positive integers, and we define the Zimin patterns  $Z_i$  inductively by  $Z_1 = v_1$  and  $Z_{i+1} = Z_i v_{i+1} Z_i$ . It turns out that, for all integers  $i, m \geq 1$  and all alphabets  $\mathcal{A}$  of cardinality  $m$ , there exists an integer  $L_{ab}(m, Z_i)$  such that all words  $\mathbf{a} \in \mathcal{A}^*$  with length at least  $L_{ab}(m, Z_i)$  contain the word  $Z_i$  in the abelian sense.

For all integers  $m \geq 4$ , Tao proves in [1] that  $L_{ab}(m, Z_i) \geq (1 + \varepsilon_m(i)) \sqrt{\mathbf{K}(m, i)}$  for all  $i \geq 1$ , where  $\varepsilon_m$  is a function such that  $\lim_{+\infty} \varepsilon_m = 0$  and  $\mathbf{K}(m, i)$  is defined as

$$\mathbf{K}(m, i) = 2 \prod_{j=1}^{i-1} \mathbf{S}(m, 2^j)^{-1},$$

$$\text{where } \mathbf{S}(m, k) = \sum_{\ell=1}^{\infty} \mathbf{T}(m, k, \ell) \text{ and } \mathbf{T}(m, k, \ell) = \frac{1}{m^{k\ell}} \sum_{i_1 + \dots + i_m = \ell} \binom{\ell}{i_1 \dots i_m}^k.$$

Yet, in order to obtain actual lower bounds on  $L_{ab}(m, Z_i)$ , it remains to evaluate the asymptotical behavior of  $\mathbf{K}(m, i)$ . We evaluate  $\mathbf{K}(m, i)$  up to a multiplicative constant that does not depend on  $m$  or  $i$ . More precisely, we prove the following inequalities, which hold for all  $m \geq 4$  and  $i \geq 1$ :

$$2 \frac{m^{2^i}}{m^{i+1}} \geq \mathbf{K}(m, i) \geq \frac{1}{21} \frac{m^{2^i}}{m^{i+1}}.$$

\*This work is supported by ERC EQualIS (308087).

†LSV, CNRS & ENS Cachan, Univ. Paris-Saclay, France

## 2 Auxiliary inequalities

Before evaluating the lower bound  $\mathbf{K}(m, i)$ , we prove a series of six inequalities that we will use subsequently. We first study the function  $f_x : y \mapsto y \ln(1 + x/y)$  for  $x, y > 0$ . An asymptotic evaluation proves that  $\lim_{y \rightarrow \infty} f_x = x$ . Furthermore, we compute that  $f_x''(y) = -\frac{x^2}{y(x+y)^2} < 0$  for  $y > 0$ . It follows that  $x > f_x(y)$  or, equivalently, that

$$(1 + x/y)^y < e^x \quad \text{for all } x, y > 0. \quad (1)$$

We perform a similar study with the function  $g : y \mapsto (y + 1/2) \ln(1 + 1/y)$  for  $y > 0$ . We find that  $\lim_{y \rightarrow \infty} g = 1$  and that  $g''(y) = \frac{1}{2y^2(y+1)^2} > 0$  for  $y > 0$ . It follows that  $g(y) > 1$  or, equivalently, that

$$(1 + 1/y)^{y+1/2} > e \quad \text{for all } y > 0. \quad (2)$$

Again, we consider the function  $h : y \mapsto 3 \ln(y) + \ln(2) - (y - 1) \ln(2\pi)$  for  $y > 0$ , as well as the real constant  $\lambda = \frac{4}{\pi^{3/2}} < \frac{3}{4}$ . We find that  $h'(y) = \frac{3}{y} - \ln(2\pi) < 0$  when  $y \geq 4$  and that  $\exp(h(4)) = \frac{16}{\pi^3} = \lambda^2$ , and it follows that

$$2y^3 \leq \lambda^2 (2\pi)^{y-1} \quad \text{for all } y \geq 4. \quad (3)$$

Similarly consider the function  $\bar{h} : y \mapsto 5 \ln(y) + \ln(2) - (y - 1) \ln(2\pi)$  for  $y > 0$ . We find that  $\bar{h}'(y) = \frac{5}{y} - \ln(2\pi) < 0$  when  $y \geq 7$  and that  $\exp(\bar{h}(7)) = \frac{7^5}{32\pi^6} < 1$ , and it follows that

$$2y^5 \leq (2\pi)^{y-1} \quad \text{for all } y \geq 7. \quad (4)$$

Then, we set  $\mathbf{Z}(x) = \sqrt{2\pi} x^{x+1/2} e^{-x}$  for all  $x \geq 0$ . We prove below that

$$\frac{(a+b)!}{a!b!} \leq \frac{\mathbf{Z}(a+b)}{\mathbf{Z}(a)\mathbf{Z}(b)} \quad \text{for all integers } a, b \geq 1. \quad (5)$$

We study the functions  $F : (a, b) \mapsto \frac{(a+b)!\mathbf{Z}(a)\mathbf{Z}(b)}{\mathbf{Z}(a+b)a!b!}$  and  $G : (a, b) \mapsto \frac{F(a+1, b)}{F(a, b)}$ . We compute that

$$\begin{aligned} G(a, b) &= \left( \frac{a+b}{a+b+1} \right)^{a+b+1/2} \left( \frac{a+1}{a} \right)^{a+3/2} \geq \bar{G}(a, b)^{2a+b+2}, \text{ where} \\ \bar{G}(a, b) &= \frac{2a+b+2}{(a+b+1/2)^{\frac{a+b+1}{a+b}} + (a+3/2)^{\frac{a}{a+1}}} \quad (\text{by geometric-harmonic inequality}) \\ &= 1 + \frac{b-1}{(2a+2b+1)(a+b+1)(a+1) + (2a+3)a(a+b)}. \end{aligned}$$

and since  $b \geq 1$ , it follows that  $G(a, b) \geq \bar{G}(a, b)^{2a+b+2} \geq 1$ , i.e. that  $F(a, b) \leq F(a+1, b)$ . Since  $F(a, b) = F(b, a)$  for all  $a, b \geq 1$ , we derive immediately that  $F(a, b) \leq F(a, b+1) \leq F(a+1, b+1)$  for all integers  $a, b \geq 1$ . Moreover, Stirling's approximation formula states that  $a! \sim \mathbf{Z}(a)$  when  $a \rightarrow +\infty$ . This proves that  $\lim_{\alpha, \beta \rightarrow +\infty} F(\alpha, \beta) = 1$ , and it follows that  $F(a, b) \leq 1$  for all  $a, b \geq 1$ , which is indeed equivalent to the inequality (5).

As a corollary, observe that, for all integers  $i_1, \dots, i_m \geq 1$  and using inequality (5), we also have

$$\frac{(i_1 + \dots + i_m)!}{i_1! \dots i_m!} = \prod_{j=2}^m \frac{(i_1 + \dots + i_j)!}{(i_1 + \dots + i_{j-1})! i_j!} \leq \prod_{j=2}^m \frac{\mathbf{Z}(i_1 + \dots + i_j)}{\mathbf{Z}(i_1 + \dots + i_{j-1}) \mathbf{Z}(i_j)},$$

from which follows our last auxiliary inequality:

$$\frac{(i_1 + \dots + i_m)!}{i_1! \dots i_m!} \leq \frac{\mathbf{Z}(i_1 + \dots + i_m)}{\mathbf{Z}(i_1) \dots \mathbf{Z}(i_m)} \quad \text{for all integers } i_1, \dots, i_m \geq 1. \quad (6)$$

### 3 Evaluating $\mathbf{K}(m, i)$

We first evaluate  $\mathbf{T}(m, k, \ell)$  when  $\ell = 1$ . Here, instead of considering a tuple of non-negative integers  $(i_1, \dots, i_m)$  that sum up to  $\ell$ , we might directly consider the unique integer  $j \in \{1, \dots, m\}$  such that  $i_j = 1$ . Moreover, for each tuple  $(i_1, \dots, i_m)$ , the multinomial coefficient  $\binom{\ell}{i_1 \dots i_m}$  is equal to 1. It follows that  $\mathbf{T}(m, k, 1) = m^{-k} \sum_{j=1}^m 1 = m^{1-k}$ , from which we derive the inequalities  $\mathbf{S}(m, k) \geq \mathbf{T}(m, k, 1) = m^{1-k}$  and

$$\mathbf{K}(m, i) \leq 2 \prod_{j=1}^{i-1} m^{2^j-1} = 2 \frac{m^{2^i}}{m^{i+1}}.$$

Then, we investigate lower bounds of  $\mathbf{K}(m, i)$ , i.e. upper bounds of  $\mathbf{T}(m, k+1, \ell)$  and of  $\mathbf{S}(m, k+1)$  when  $m \geq 4$  and  $k \geq 1$ . Consider some integer  $\ell \geq 1$ , and let us write  $\ell = am + b$ , with  $a \geq 0$  and  $1 \leq b \leq m$ . In addition, let us set

$$\mathbf{V}(m, a, b) = \max_{i_1 + \dots + i_m = \ell} \binom{\ell}{i_1 \dots i_m} \text{ and } \mathbf{U}(m, a, b) = \frac{1}{m^\ell} \mathbf{V}(m, a, b).$$

We first observe that

$$\begin{aligned} \mathbf{T}(m, k+1, \ell) &= \frac{1}{m^{(k+1)\ell}} \sum_{i_1 + \dots + i_m = \ell} \binom{\ell}{i_1 \dots i_m}^{k+1} \\ &\leq \frac{1}{m^{(k+1)\ell}} \sum_{i_1 + \dots + i_m = \ell} \binom{\ell}{i_1 \dots i_m} \mathbf{V}(m, a, b)^k \\ &\leq \mathbf{U}(m, a, b)^k \mathbf{T}(m, 1, \ell) \\ &\leq \mathbf{U}(m, a, b)^k. \end{aligned} \quad (\text{by Newton multinomial identity})$$

Since the inequality  $(x+1)!(y-1)! \geq x!y!$  holds for all integers  $x \geq y$ , it also follows that

$$\mathbf{U}(m, a, b) = \frac{(am+b)!}{m^{am+b}(a+1)!b!a^{m-b}}.$$

We compute immediately that  $\mathbf{U}(m, 0, b) = m^{-b}b! = m^{-1}$  if  $b = 1$ , and that  $\mathbf{U}(m, 0, b) \leq 2m^{-2}$  if  $2 \leq b \leq m$ . When  $a \geq 1$ , we further compute that

$$\begin{aligned} \mathbf{U}(m, a, b) &\leq \frac{\mathbf{Z}(am+b)}{m^{am+b}\mathbf{Z}(a+1)^b\mathbf{Z}(a)^{m-b}} && (\text{using inequality (6)}) \\ &\leq \frac{\sqrt{am+b}}{(2\pi)^{(m-1)/2}a^{m/2}} \frac{(1+b/am)^{am}(1+b/am)^b}{(1+1/a)^{(a+3/2)b}} \\ &\leq \frac{\sqrt{2am}}{(2\pi)^{(m-1)/2}a^{m/2}} \frac{(1+b/am)^{am}(1+1/a)^b}{(1+1/a)^{(a+3/2)b}} && (\text{since } b \leq m \leq am) \\ &\leq \frac{\sqrt{2m}}{(2a\pi)^{(m-1)/2}} \frac{(1+b/am)^{am}}{(1+1/a)^{(a+1/2)b}} \\ &\leq \frac{\sqrt{2m}}{(2a\pi)^{(m-1)/2}} \frac{e^b}{(1+1/a)^{(a+1/2)b}} && (\text{using inequality (1)}) \\ &\leq \frac{\sqrt{2m}}{(2a\pi)^{(m-1)/2}} \left( \frac{e}{(1+1/a)^{a+1/2}} \right)^b \\ &\leq \frac{\sqrt{2m}}{(2a\pi)^{(m-1)/2}}. && (\text{using inequality (2)}) \end{aligned}$$

Consequently, since  $k \geq 1$ , we find that

$$\begin{aligned} \mathbf{S}(m, k+1) &\leq \sum_{a=0}^{\infty} \sum_{b=1}^m \mathbf{U}(m, a, b)^k = \mathbf{U}(m, 0, 1)^k + \sum_{b=2}^m \mathbf{U}(m, 0, b)^k + \sum_{a=1}^{\infty} \sum_{b=1}^m \mathbf{U}(m, a, b)^k \\ &\leq \frac{1}{m^k} + \frac{2(m-1)}{m^{2k}} + m \sum_{a=1}^{\infty} \left( \frac{\sqrt{2m}}{(2a\pi)^{(m-1)/2}} \right)^k \\ &\leq \frac{1}{m^k} + \frac{2m}{m^{2k}} + m \frac{(2m)^{k/2}}{(2\pi)^{k(m-1)/2}} \zeta(k(m-1)/2). \end{aligned}$$

If we set

$$\mathbf{P}(m, k) = 1 + 2m^{1-k} + m^{k+1} \frac{(2m)^{k/2}}{(2\pi)^{k(m-1)/2}} \zeta(k(m-1)/2),$$

then it follows that  $\mathbf{S}(m, k+1) \leq \frac{1}{m^k} \mathbf{P}(m, k)$  and therefore that

$$\mathbf{K}(m, i) \geq 2 \prod_{j=1}^{i-1} \frac{m^{2^j-1}}{\mathbf{P}(m, 2^j-1)} \geq \frac{2}{\mathbf{P}_{\infty}(m)} \frac{m^{2^i}}{m^{i+1}},$$

where  $\mathbf{P}_{\infty}(m)$  is the infinite product  $\prod_{j=1}^{\infty} \mathbf{P}(m, 2^j-1)$ . It remains to prove that  $\mathbf{P}_{\infty}(m) \leq 42$ .

We first assume that  $7 \leq m$ . For  $k \geq 1$ , we compute that

$$\begin{aligned} \mathbf{P}(m, k) &= 1 + 2m^{1-k} + m^{k+1} \frac{(2m)^{k/2}}{(2\pi)^{k(m-1)/2}} \zeta(k(m-1)/2) \\ &= 1 + 2m^{1-k} + m^{1-k} \left( \frac{2m^5}{(2\pi)^{m-1}} \right)^{k/2} \zeta(k(m-1)/2) \\ &\leq 1 + 2m^{1-k} + m^{1-k} \zeta(k(m-1)/2) \quad (\text{using inequality (4)}) \\ &\leq 1 + 4m^{1-k} \quad (\text{since } \zeta(k(m-1)/2) \leq \zeta(3) \leq 2) \\ &\leq \exp(4m^{1-k}), \quad (\text{since } 1+x \leq \exp(x) \text{ for all } x \in \mathbb{R}) \end{aligned}$$

from which we deduce that  $\mathbf{P}_{\infty}(m, 1) \leq 5$ , whence

$$\begin{aligned} \ln(\mathbf{P}_{\infty}(m)) &= \sum_{j=1}^{\infty} \ln(\mathbf{P}(m, 2^j-1)) \leq \ln(\mathbf{P}(m, 1)) + 4 \sum_{j=2}^{\infty} m^{2-2^j} \leq \ln(5) + 4 \sum_{j=0}^{\infty} m^{-2-j} \\ &\leq \ln(5) + \frac{4}{m(m-1)} \leq \ln(5) + \frac{2}{21} \leq \ln(42). \quad (\text{since } 7 \leq m) \end{aligned}$$

Then, we assume that  $4 \leq m \leq 6$ . Again, for  $k \geq 1$ , we compute that

$$\begin{aligned} \mathbf{P}(m, k) &= 1 + 2m^{1-k} + m^{k+1} \frac{(2m)^{k/2}}{(2\pi)^{k(m-1)/2}} \zeta(k(m-1)/2) \\ &= 1 + 2m^{1-k} + m \left( \frac{2m^3}{(2\pi)^{m-1}} \right)^{k/2} \zeta(k(m-1)/2) \\ &\leq 1 + 2m^{1-k} + m \zeta(k(m-1)/2) \lambda^k \quad (\text{using inequality (3)}) \\ &\leq 1 + 2m^{1-k} + 3m \lambda^k \quad (\zeta(k(m-1)/2) \leq \zeta(3/2) \leq 3) \\ &\leq 1 + 5m \lambda^k \quad (\text{since } m^{-1} \leq \frac{1}{4} \leq \lambda) \\ &\leq \exp(5m \lambda^k) \quad (\text{since } 1+x \leq \exp(x) \text{ for all } x \in \mathbb{R}) \end{aligned}$$

Furthermore, explicit computations in each of the cases  $m = 4$ ,  $m = 5$  and  $m = 6$  indicate that  $\prod_{j=1}^4 \mathbf{P}(m, 2^j - 1) \leq 41$ . Hence, we conclude that

$$\begin{aligned}
\ln(\mathbf{P}_\infty(m)) &\leq \ln(41) + \sum_{j=5}^{\infty} \ln(\mathbf{P}(m, 2^j - 1)) \\
&\leq \ln(41) + 5m \sum_{j=5}^{\infty} \lambda^{2^j - 1} \leq \ln(41) + 5m \sum_{j=0}^{\infty} \lambda^{31+j} \\
&\leq \ln(41) + \frac{5m\lambda^{31}}{1-\lambda} \leq \ln(41) + \frac{30 \times 3^{31}}{4^{30}} \leq \ln(42). \quad (\text{since } m \leq 6 \text{ and } \lambda < \frac{3}{4})
\end{aligned}$$

## References

- [1] J. Tao. Pattern occurrence statistics and applications to the Ramsey theory of unavoidable patterns. *ArXiv e-prints*, June 2014.